Enhancing Information Management in US DOE's FUSRAP – 23254

Tony Pace*, Scott Raynes, Daniel List, and Darina Castillo **
* RSI EnTech, LLC
** US Department of Energy, Office of Legacy Management

ABSTRACT

The Formerly Utilized Sites Remedial Action Program (FUSRAP) was created in the mid-1970s to clean up radiological contamination resulting from early development of nuclear weapons. The US Department of Energy (DOE) was responsible for FUSRAP until October 1997, when Congress transferred administration and execution of FUSRAP site cleanup to the US Army Corps of Engineers (USACE). DOE's Office of Legacy Management (LM) determines the eligibility of a site for referral to USACE. USACE determines whether the site is included in FUSRAP for cleanup, conducts remedial actions, and transfers sites back to LM for long-term stewardship (LTS).

As with all information-driven programs, LM's FUSRAP acquires and creates an abundance of documents, data, photographs, and other records in various formats from multiple sources. To modernize information management, LM is implementing applications and storage systems that are flexible and scalable to the various programs. This effort will reduce duplication, ensure efficiency in data management, and increase accessibility to information for programmatic decision-making. Primary areas of focus include geospatial and environmental data, photographic and video media, and data management.

Geospatial and environmental data are stored in a variety of internal databases to support LM's LTS mission. Efforts were made to provide an interface for LM and Legacy Management Support team members to access programmatic and site-specific information and to streamline the creation and updating of executive briefing materials. Based on this effort, a geographic information system-based interactive web interface has been developed to retrieve site-specific information and data from one central source, which provides an intuitive user experience. This web interface allows everyday users to access authoritative data quickly and enables future analysis of performance and process via graphics, widgets, maps, and custom tools.

A FUSRAP media library was implemented as part of a larger organizational effort to establish governance over LM's photos and videos including metadata standards, resolution and format guidance, naming convention best practices, and so on. The FUSRAP team's digital photos are now stored in the media library on the internal LM website. The tool allows for adding metadata to content and supports searching via metadata and browsing via thumbnails. When these media are no longer actively used, they are archived to LM's official Enterprise Content Management system.

The Considered Sites Database (CSD), compiled by a predecessor agency, includes relevant eligibility and remediation information for sites considered under FUSRAP and is available to the public via the LM CSD webpage. The Master Site List (MSL) was developed by LM to summarize data about all sites including sites referred to but rejected as ineligible under FUSRAP. The MSL contains similar metadata fields, but the two sources do not fully align. As part of a larger authoritative data source initiative, FUSRAP is considering options for consolidating the CSD and MSL to reduce redundancy and better align the two datasets while serving different purposes. The consolidated database would feed several systems, including the public website and web interface, thereby ensuring consistency, and reducing manual comparisons and system discrepancies.

LM envisions future enhancements to its information management systems to enhance the LTS of its sites. The LM FUSRAP team will continue to adapt to its increasing scope and improving technology to support its mission to protect human health and the environment.

INTRODUCTION

On December 15, 2003, the Office of Legacy Management (LM) was formally established as a new US Department of Energy (DOE) office to provide a long-term, sustainable solution to the legacy of the Cold War. LM "preserves, protects, and makes accessible legacy records and information" as part of its core functions. LM sites fall under a variety of regulatory and functional categories, one of which is the Formerly Utilized Sites Remedial Action Program (FUSRAP). In 1974, FUSRAP was established to remediate sites where radioactive contamination remained from Manhattan Engineer District (MED) and early U.S. Atomic Energy Commission (AEC) operations. Since its inception, the FUSRAP Program relied heavily on historic information to identify and remediate sites. Today, FUSRAP staff and stakeholders still rely on historic information for almost 700 sites and additional information generated since then to provide long-term stewardship (LTS) to 34 sites and prepare for the transfer of 21 sites that are in active remediation. During site transfers, LM receives information including records, environmental data, and other forms of data generated by the transferring agency. As sites transferring to LM become larger and more complex, LM anticipates that the volume of information that transfers with each of the 21 sites will increase. Additionally, LM continues to generate its own records and other types of information on many sites within the program. In order to manage the extensive collection of information, LM developed tools and management strategies to aid site management and LTS responsibilities. These tools will reduce duplication, ensure efficiency in information management, and increase accessibility to information for programmatic decision-making. Primary areas of focus include geospatial and environmental data, photographic and video media, and data management. The establishment of authoritative data sources is critical to ensuring that decisions are made with recognized and confirmed data that are validated, reliable, and accurate.

DESCRIPTION

LM currently manages more than 3,140 cubic meters of physical records, more than 4.2 million electronic records and finding aids, and large volumes of data to support its mission. In addition to generating its own content, LM receives records and information for transitioning sites from other agencies. Much of LM's collection of information includes FUSRAP records that were created as early as the 1950s to support early MED and AEC activities, such as contracts with private entities, to support the Manhattan Project and Cold War missions. Activities were conducted by multiple private and public entities and administered from multiple locations. Consequently, records were stored across the entire complex in government, corporate, and academic collections. These sites continue to generate additional information in a variety of formats. Historical information in LM's collections are relied upon today to make LTS decisions and to respond to stakeholder inquires related to human health and the environment. While improving accessibility LM has also sought to improve reliability and accuracy of the information it manages.

During this multiyear effort LM focused on the accessibility of environmental data, photographs, and videos as well as historical records by developing a variety of tools on application platforms that are readily utilized in the LM environment to ensure flexibility and scalability. The conversion of information from historical formats, such as scanned PDF files or Microsoft Word™ documents, supports the overall goals of the program.[a] For example, environmental data historically found in old reports were moved into environmental databases and ultimately to Esri™ Story Maps, historic photographs in boxes were made accessible through a SharePoint™ based photo repository, and information that was previously generated from multiple data sources is being consolidated where possible and updated to reflect additional information gathered since the start of the FUSRAP Program.[bc]

[a] *Microsoft Word™ is a trademark of the Microsoft Corporation in the United States and/or other countries*.
[b] *Esri™ is a trademark of the Esri Corporation in the United States and/or other countries*.
[c] *SharePoint™ is a trademark of the Microsoft Corporation in the United States and/or other countries.*

In conjunction with the information management tools, updated records and information strategies were generated to align with updated records and information management practices and the management strategy of the FUSRAP Program. Data Governance best practices were essential to the design of the tools and the information management guidelines that were used to maintain the tools. The FUSRAP Program now has several site managers and an extensive list of functional support team members that support LTS at the sites. To maintain the developed tools and ensure accuracy and reliability of the information generated, procedures and plans were developed, tutorial sessions were conducted, and contractual requirements were updated.

DISCUSSION

The tools that LM has generated required cross-functional teams and the involvement of disparate groups with common interests. Cross-functional teams often collaborate on projects to ensure that the tools being developed have input from the various stakeholders and provide the greatest usability. The geospatial and environmental data, records management, and site teams on both federal and contractor levels were critical in the development of new tools and management strategies that are described in this paper.

Geospatial and Environmental Data

FUSRAP project personnel and Environmental and Spatial Data Management (ESDM) staff formed the FUSRAP working group comprised of members from LM and Legacy Management Support (LMS) staff. Together, they worked to realize the benefits of a web geographic information system and allow internal users the ability to access authoritative data and other programmatic and site-specific information through an Esri™ story map-driven application within the LM GeoPortal. The interactive FUSRAP Web Interface has been a three-year, three-phase project between FUSRAP and ESDM teams that was initiated in federal fiscal year 2020.

With this three-phased approach, the FUSRAP and ESDM team members mapped out a course using the Responsible, Accountable, Consulted, Informed (RACI) model and worked to bring together the elements that form the current version of the FUSRAP Web Interface. A project charter was used to define the scope and goals of each phase, maintain the focus of the team, and establish clear and attainable milestones.

In Phase 1, team members developed the scope to initiate the establishment of the RACI model, the roles, and responsibilities of the team members for the selection of data, application design and development, and the operation and maintenance of the application. Once development of the initial project requirements was completed, a technical solution was found, and a prototype was developed. This prototype, which utilized Esri™ story maps, was intended to collate and communicate key program and site-specific information that is currently maintained in separate work products, reports, and summaries. The story map concept provides a mechanism for merging data, text, and multimedia products in an interactive interface. Team members began determining the key elements within the application that should be included. Many elements were documents that are currently used by LM/LMS management to assess current site status and conditions or simply to gain familiarity with the site. The executive briefing materials served as the model for inclusion of program- and site-specific content in this application. Descriptions of the site and its history, status, requirements, and so on, were compiled from these materials and other summary documents and existing resources.

At the time of development, there was no known repository for media, and the ESDM team was undertaking a data reorganization and migration to better support LM across all sites and programs. Much of the graphic content presented in the FUSRAP Web Interface copied from existing resources, saved to SharePoint™, and repurposed in the application on a case-by-case basis. A catalog of briefing materials

was not developed given the distributed nature of source information; however, to achieve the same goal, a SharePoint™ library was organized and used to stage content for inclusion in the application and cited source documents when appropriate. In turn, this propagated the media library, another information management enhancement, which is discussed in detail in the next section

Links to key documents and external resources were incorporated into the application as well. Ongoing maintenance after project closure is required and addressed in an operations and guidance document developed in Phase 2 and implemented in Phase 3. A thorough review of the prototype solicited feedback so that Phase 2 development would be more informed.

The Phase 1 prototype was developed, tested, and published to the FUSRAP Working Group within the LM GeoPortal with the goal of addressing all Phase 1 commitments and proving, in concept, that all stated needs and requirements could be met within the Esri™ story maps environment using established Esri™-provided templates. A subset of eight sites was used, representing both "Active" and "Completed" status and involving all LMS site leads and LM site managers. Active sites include those currently managed or being remediated by US Army Corps of Engineers (USACE) and completed sites have been remediated and transitioned to LM for LTS. After demonstrating interface capabilities in Phase 1, including the ability to filter, search, and link to a variety of content, feedback was received from both LM and LMS stakeholders to be incorporated during Phase 2 planning and development.

During Phase 2, the production team began developing the web interface and using the feedback and lessons learned to make improvements and add content while building on the foundations established in Phase 1. This phase also recognized the need to better document and access information on "Considered" sites. Site content from this prototype was reorganized into new tabs and headers to better reflect the structure of the executive briefing materials. LMS site leads then proceeded to compile content for the remaining Active sites in Microsoft Word™ files designed to mimic the agreed-upon structure of Active site story maps. A similar process was used for creating the Completed sites story maps. A new program landing page was developed using Esri™ dashboards during Phase 2 which offered a program overview and provided access to individual site pages. A task was performed to map the locations of these Considered sites to support the development of a Considered Sites webpage in Phase 3.

While Phase 2 focused on developing the application and populating it with selected content, the scope for Phase 3 in the project charter was about making enhancements and aligning this tool with the organization's geospatial strategy and the development efforts outlined in the *Enterprise Geospatial Strategy Implementation Plan*. [1] During Phase 3, an operations and maintenance (O&M) plan was developed to document the conceptual layout, as well as processes and procedures of O&M activities for the web interface. Concurrent to the development of O&M activities, which included training site subject matter experts to edit in the story map environment, ESDM developed and refined a Considered Sites webpage within the Esri™ dashboard template. This dashboard (Figure 1) provides a detailed overview of Considered sites, as well as filtering and search tools to visualize the locations of these sites.

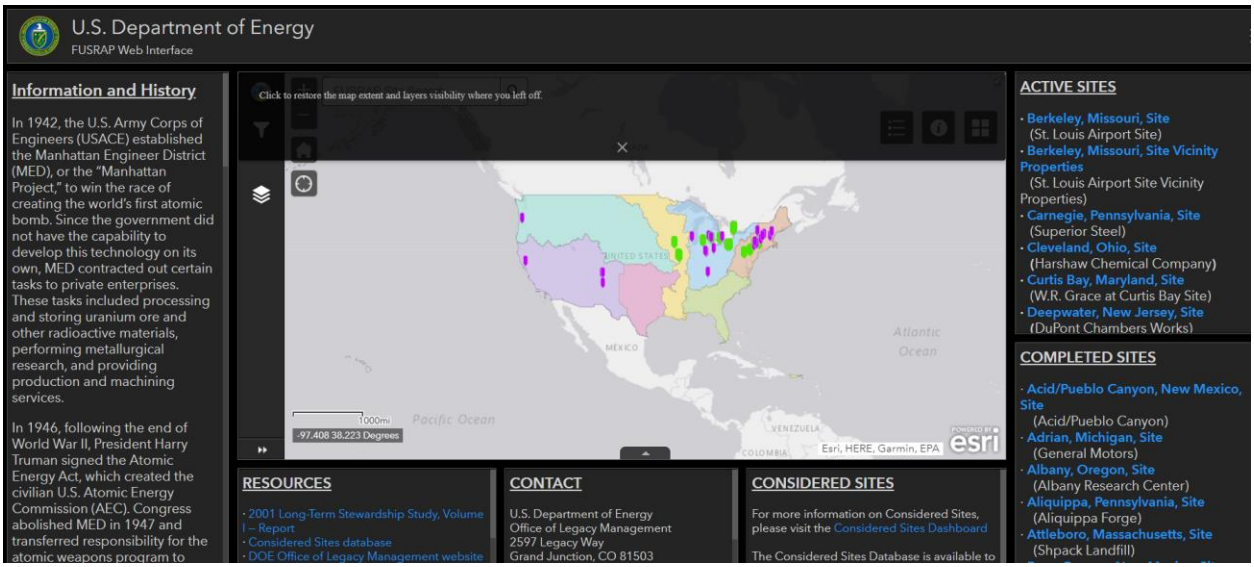

Fig. 1. Interactive FUSRAP Web Interface.

As sites are visited or transitioned from USACE to LM, additional photographs, graphics, and information will become available for use and included in the media library.

Media Library

Also in 2020, LM initiated a project to identify photos of FUSRAP sites and activities to consolidate them into LM's SharePoint™ environment. The staff identified about 3000 photos and captured them into SharePoint™ libraries organized by site type (Active, Completed, Ineligible) and site name for ease of use. However, certain limitations were noted including the inability to preview the photos without opening them and photos from events that spanned multiple sites not always being filed within the appropriate site folder.

In 2022, the team was tasked with maintaining and enhancing the photo management tool to include developing processes to support photo management tasks. During the review, the team identified a method for consolidating the photos from multiple SharePoint™ document libraries into a single SharePoint™ media library, which yielded several immediate benefits. Autogenerated photo thumbnails could be browsed or viewed alongside tailored metadata fields. Custom lists using choice fields were used to ensure that consistent site naming and collections of photos could be grouped under a single event or title. In addition, detailed descriptive metadata, sometimes pulled from file names, could be added and searched. Filters could be used to view photos by site, event, date, and so on. Metadata fields were aligned with the National Archives and Records Administration (NARA) guidance, as well as LM's Enterprise Content Management (ECM) system to allow for future archiving. As part of the migration, limited quality control reviews were conducted.

Shortly after the media library was introduced, the team was tasked with a project to perform a self-assessment of the photos and how they were being managed. This self-assessment, to be performed biannually, allowed for a more in-depth review of the photos and metadata. Numerous duplicate photos were identified and eliminated. Also, event titles, descriptions, and dates were populated or expanded, to the extent possible. New photos were identified in the FUSRAP network file shares and moved into the library, and a standard naming convention was developed and applied to the new photos. In addition to photos, hyperlinks were added to FUSRAP site photos residing outside of the LM environment, such as those at NARA and the Library of Congress. Additional sources of FUSRAP site inspection photos were

identified within the LM network and, rather than duplicating them, links were created to point to those collections that could not be moved for numerous reasons. The final task of the self-assessment involved updating the job aid, which was initially developed during the migration, to reflect improved processes; the self-assessment process was documented to make the workflow easier for the next performance period.

Over the course of two years, the FUSRAP team developed a photo management tool (Figure 2) that allows for the browsing and searching of over 3300 photos with more than 70 links to an additional 6800 photos.

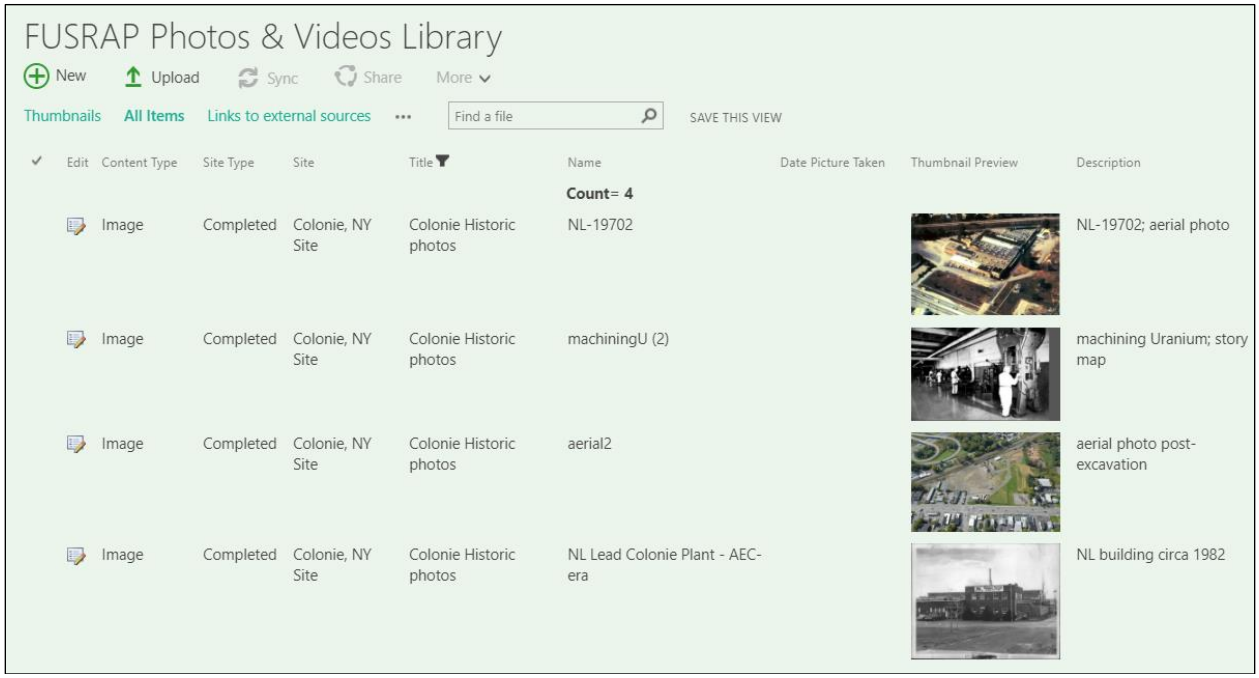

Fig. 2. Media Library with Metadata.

Considered Sites Database (CSD) and Master Site List (MSL)

The FUSRAP team maintains separate data sources for information accessible to the public and those used for internal support functions. The two sources contain similar data, but each evolved independently and serve slightly different purposes. The CSD, compiled by a predecessor agency, includes relevant eligibility and remediation information for sites considered under FUSRAP, and is available to the public via the LM CSD webpage. The MSL was initially developed by LM in 1995 around several source lists and was used as an internal tool to respond to *The Wall Street Journal*'s information [2] about former weapons sites. The MSL has evolved over the years and now summarizes data from additional sources about all sites including those referred to but rejected as ineligible under FUSRAP. The MSL contains similar metadata fields, but the two sources do not fully align. The CSD is viewed as a static data source, whereas the MSL is updated annually. A recent exercise to compare the two sources identified numerous inconsistencies and it was complicated by differing and overlapping metadata fields.

As a result of the comparison exercise, the most significant discrepancy noted was in the site name. Many sites have multiple names dependent on the context of the audience or original data source. This has led to individual sites having upwards of five or more names, creating the need to maintain the historical name of each site as it relates to individual groups, news, or government agencies. There were also instances

where the site designation (e.g., NY01, NY-01, NY 01, or NY 0-01) was used as an identifier and applied to multiple sites with a common connection, such as the name of a university, corporate affiliation, or locale. Some sites bore the same name and were in entirely different states and, in some cases, just down the street. Each site would require a comparison of known facts, such as a street address or set of coordinates, to tie it to one location and its subsequent reconciliation.

The staff is working to review the discrepancies and determine which changes should be applied, considering the differing intended consumers of the data. A cascading effect occurs when systems with similar data must be updated.

"Data Governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods." [3] Maintaining integrity, usability, and consistency of data are the focal points to ensuring a comprehensive data management program. With these points in mind, the team established these concepts for data quality management.

Data Governance

LM recognizes the need to continually evaluate processes and best practices and to work smarter rather than harder. An increased awareness and focus on Data Governance and the need for authoritative data sources has emerged. A Data Governance policy [4] was issued in January 2022 establishing a Data Governance Council and several working groups to develop standards and policies in the areas of environmental data, geospatial data, and digital media (e.g., photos and videos).

An authoritative data source is defined as a repository or system that contains information considered to be the primary and most reliable source for this information. Authoritative data will be visible, accessible, understandable, linked, trusted, and secure. If two or more systems have mismatched or conflicting data, the information within the authoritative data source is considered to be the most accurate. [5]

As part of a Data Governance initiative and the establishment of authoritative data sources, the staff is discussing options for optimizing the CSD and MSL to reduce redundancy and better align the two datasets while recognizing that they serve different purposes. Authoritative data sources would feed several systems, including the LM public website and the FUSRAP Web Interface, thereby ensuring consistency and reducing the need for manual comparisons and potential for system discrepancies.

Legacy Data Cleanup

LM's data management has evolved over the years as new technology becomes available. Network shared folders (also called file shares) were widely used for storing working files and sharing information within projects and teams. As record copies were submitted and captured in the Electronic Recordkeeping System (ERKS), it was common practice to retain reference copies in file shares for convenience. SharePoint™ was introduced for improved collaboration and governance but was not consistently adopted.

In 2019, LM migrated from an ERKS to an ECM system. While ERKS provided a repository for official records with limited access, the ECM provides improved information management functionality including access for all LM and LMS staff, and the filing of their own content, which was previously achieved by a centralized records management staff. General access eliminates previously siloed information, reducing the need for redundancy and retention of reference copies. The ECM serves as LM's archival system and records repository. As part of the larger Data Governance initiative, LM is shifting from file shares to maintained working files in SharePoint™ and archived records and inactive content in ECM. This shift involves decades of legacy files that need to be appraised for record content and considered for continued retention (Figure 3).

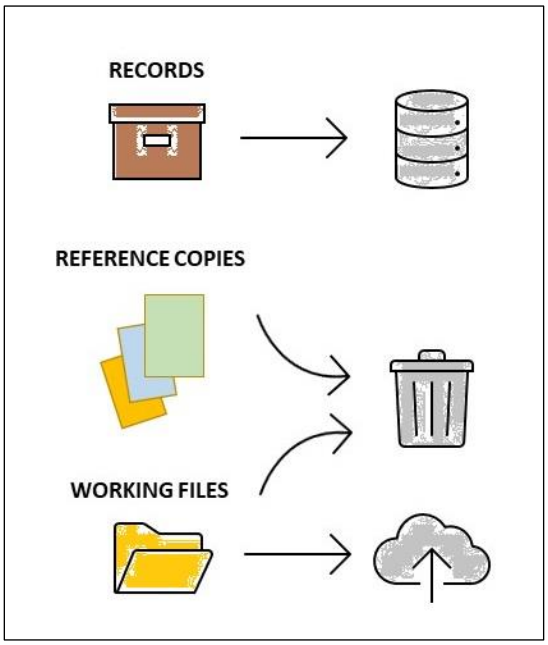

Fig. 3. Legacy Data Cleanup.

CONCLUSIONS

Given the recent implementation, the FUSRAP Web Interface has already proven to be a highly successful and useful tool. The staff continues to find new uses including changing references of source documents into hyperlinks and refining existing data. The tool is expected to evolve as new needs emerge and potential uses are identified.

Events and LTS activities continue to generate photos, and new external sources of FUSRAP photos continue to be identified. Opportunities for process improvements for photo and video management will be evaluated for potential consolidation, de-duplication, and standardization as policies are issued and formally recognized repositories are defined.

The FUSRAP team has established a tracking log for potential updates to the CSD and MSL. To date, more than 250 potential changes were identified and queued for review and more than 50 (20%) have been approved and applied. The staff plans a thorough review of the CSD and MSL to validate the information, tie it back to source documents to the extent possible, and ensure consistency in formatting, phrasing, and terminology.

As LM's Data Governance initiative matures, policies will be issued or updated, and working groups will continue to focus on areas for the improved management of LM's data assets. The Data Governance Council will evaluate working group-developed policies to ensure consistency within the LM enterprise.

The legacy data cleanup initiative continues as a long-term effort to achieve improved governance over LM's network file shares. To date, cleanup efforts have yielded a 43% reduction in the number of files stored in the FUSRAP team's file shares, through the deletion of working files and reference copies, capture of records, and migration of files to SharePoint™. A shared drive policy is being drafted that will clarify appropriate use of network file shares with a focus on reduction of network storage and improved processes for dispositioning working files and elimination of reference copies. Legacy data cleanup projects are ongoing.

Information management continues to evolve as data silos are identified, policies are implemented, and technology evolves. LM continues to evaluate best practices and work toward authoritative data sources to ensure that decisions are made based on accurate information from recognized data sources. LM continues to identify cross-functional activities where the potential for redundancy exists. This recognition of overlapping needs will continue to guide future Data Governance initiatives for the foreseeable future.

REFERENCES

1. DOE (U.S. Department of Energy), 2019. Enterprise Geospatial Implementation Plan, prepared by Woolpert Inc. for the Office of Legacy Management, December 16.
2. WSJ (Wall Street Journal), 2013. "Waste Lands," https://www.wsj.com/graphics/waste-lands/, accessed October 2, 2022.
3. The Data Governance Institute, 2022. "Definitions of Data Governance," https://datagovernance.com/the-data-governance-basics/definitions-of-data-governance/, accessed October 3, 2022.
4. DOE (U.S. Department of Energy), 2022. *LMS Data Governance Policy*, LMS/POL/S36991.
5. DOE (U.S. Department of Energy), 2022. *Draft Authoritative Data Source Policy*, LMS/POL/41687, continually updated, prepared by the LMS contractor for the U.S. Department of Energy Office of Legacy Management.